Gene-expression profiling discriminates between cerebral malaria (CM)-susceptible mice and CM-resistant mice.

Nicolas F. Delahaye, Nicolas Coltel, Denis Puthier, Laurence Flori, Rémi Houlgatte, Fuad A. Iraqi, Catherine Nguyen, Georges E. Grau, Pascal Rihet

Gene expression profiling discriminates cerebral malaria-susceptible mice and cerebral malaria-resistant mice

Short title: Gene expression profiles in mouse malaria

Nicolas F. Delahaye[1], Nicolas Coltel[2], Denis Puthier[3], Laurence Flori[1,a], Rémi Houlgatte[3], Fuad A.Iraqi[4], Catherine Nguyen[3], Georges E. Grau[2] and Pascal Rihet[1,3]

[1]Université de la Méditerranée-IFR48, Laboratoire de Pharmacogénétique des maladies parasitaires-
EA864, Marseille, France
[2]Université de la Méditerranée-IFR48, CNRS-UMR 6020-Immunopathology group, Marseille, France
[3]INSERM ERM 206-TAGC, Marseille, France
[4]International Livestock Research Institute (ILRI), Nairobi, Kenya

Word counts: 146 (abstract) and 3 497 (text)

Financial support: French Ministry of Research and Technology (PAL+ Program); Fondation pour la Recherche Médicale; PACA Conseil Régional; Conseil Général des Bouches du Rhône. LF was supported by a studentship from the Fondation pour la Recherche Médicale. NFD and NC are supported by a studentship from the French Ministry of Research and Technology.
The authors do not have a commercial or other association that might pose a conflict of interest.

Reprints or correspondence: Dr. Pascal Rihet, Laboratoire de Pharmacogénétique des maladies parasitaires-EA864, Faculté de Pharmacie, Université de la Méditerranée, IFR 48, 27 Bd Jean Moulin 13385 Marseille Cedex 5, France. Telephone/Fax: +33 4 91 80 36 74.
E-mail adress: rihet@luminy.univ-mrs.fr
[a] Present affiliation: Institut National de la Recherche Agronomique (INRA), Jouy-en-Josas, France.

Abstract

The development of mouse cerebral malaria (CM) upon *Plasmodium berghei* ANKA infection is under genetic control. Brain gene expression patterns were investigated in well-defined genetically CM-resistant (BALB/c and DBA/2) and CM-susceptible (C57BL/6 and CBA/J) mouse strains by using cDNA microarrays. By combining transcriptional profiling with rigorous statistical methods and cluster analysis, we identified a set of 69 genes that perfectly discriminated between mouse strains and between CM-resistant and CM-susceptible mice. The analysis of Gene Ontology terms revealed that the genes clustered and related to susceptibility to CM preferentially belonged to some biological process classes, such as those pertaining to immune responses. Using a false discovery rate of 5% and Welch t-test, we identified 31 genes with consistent differential expression between CM-resistant and CM-susceptible mice. These data indicate that microarray analysis may be useful to identify candidate genes potentially responsible for resistance or susceptibility to CM.

Key words: cerebral malaria, microarrays, mouse, brain, resistance, statistics, immunopathology

Introduction

Malarial disease is a major public health problem in the developing countries. Cerebral malaria (CM) is one of the most severe complications and a major cause of death [1,2]. A large body of evidence indicates that malaria infection and disease have heritable components, both in the human disease and in murine models. In the search for human genetic factors involved in malaria pathology or protection, several genes have been associated with resistance to severe malaria, and genes controlling parasitemia and mild malaria have been mapped [3-6]. The role of some genes has been demonstrated in mouse models and new malaria resistance genes that have been located on mouse genome remain to be identified [1,7].

To investigate the pathogenesis of CM, several animal models have been established, including mice infected by *Plasmodium berghei* ANKA (PbA) [7,8]. Although experimental CM models cannot reproduce all the features of human CM, several observations in mouse CM have been extended and confirmed in human disease. These include behavioural changes, histopathological features in brain and retina, biochemical changes in brain, blood-brain barrier function, immunological responses, and expression of molecules in brain and retina [1].

Inbred mouse strains markedly differ in their susceptibility to *Plasmodium* infection and disease. In the murine model of PbA infection, BALB/c and DBA/2 mice are resistant to CM (CM-R) while C57BL/6 and CBA/J mice are susceptible to the neurological syndrome (CM-S). Our working hypothesis was that CM resistance/susceptibility differences may be due to genetic differences influencing gene expression. In particular, since CM is associated with an abnormal increase in the proinflammatory immune responses, it is conceivable that genes involved in inflammation are up-regulated.

Global expression analysis of 5053 genes was employed to identify differences in the brain expression between CM-R mice (BALB/c and DBA/2 mice) and CM-S mice (C57BL/6 and CBA/J).

We report here that CM-R mice and CM-S mice infected by PbA differ in their gene expression, the gene expression profile being discriminant of resistance or susceptibility to CM.

Materials and Methods

Mouse Strains and Phenotyping

Six to 8 weeks old BALB/c, DBA/2 (CM-R strains), C57BL/6 and CBA/J (CM-S strains) female, were obtained from IFFA CREDO (Ch. River Lab, France) and kept in our facilities. Five from each strain were infected by i.p. injection of *P. berghei* ANKA. Parasitemia was monitored daily on blood smear. The CM-S mice developed a neurological syndrome which occurred 6 to 7 days after parasite inoculation with a cumulative mortality of 100 %. The CM-R mice die during the 3rd or the 4th week of infection, with severe anaemia and hyperparasitemia [7].

Organ Sampling and Histology

Brains were taken from CM-R and CM-S mice when the CM-S mice develop CM. Brains were completely removed and cut in two parts: one part was frozen in RNALater (Qiagen, TM) until RNA analysis, and the other one was embedded in Tissue Tek (Leica), snap frozen in liquid nitrogen and kept at -80°C until histopathological examination of cryosections. Histology consisted in specific antibody immunohistochemistry for VCAM-1 (Clone 429, BD Pharmingen) and ISCBP1 (C19, TEBU Bio). Afterwards, slides were photographed and protein expression was quantified using LUCIA software (Nikon) [8].

RNA Isolation and cDNA Microarray Hybridizations

Total RNA from brains was extracted using TRIzol reagent (Gibco-BRL, Life Technologies). The quality of RNA was confirmed on a formaldehyde agarose gel, and the concentration of RNA was determined by reading absorbance at 260/280 nm. Each mRNA sample extracted from an individual sample was run on a single microarray. All microarray procedures were done at our microarray core facility (http://tagc.univ-mrs.fr/). cDNA microarrays were designed and

prepared as described [9]. The microarrays used in this study were composed of 5053 sequences. PCR amplification was performed as previously described [9], and PCR products were spotted onto nylon membranes (Hybond-N+, Amersham) with a GMS-417 arrayer (Affymetrix, Santa Clara, CA). About 20% of the genes included in this clone set are represented by two or more different cDNA clones, providing internal controls to assess the reproducibility of gene expression measurements. Microarrays were hybridized with ^{33}P-labelled probes, first with an oligonucleotide sequence common to all spotted PCR products (5'TCACACAGGAAACAGCTATGAC-3'), then after stripping, with complex probes made from 5 µg of retrotranscribed total RNA. Probe preparations, hybridizations and washes were carried out as described previously [9]. After 48h hybridization, arrays were scanned with a FUJI BAS5000 machine at 25 µm resolution.

Microarray Data Analysis

Hybridization signals were quantified using ArrayGauge software (Fuji Ltd, Tokyo, Japan). All images were carefully inspected, and spots with overestimated intensities due to neighborhood effects were manually excluded. The data from 1 sample (CBA/J mouse) indicated a lower hybridization quality, so the sample was omitted from subsequent analyses. The data were filtered such that only spots with intensities that were two times greater than the median background in either microarray were used in the analysis, and the signal intensities were then corrected to take into account the amount of spotted DNA and the variability of experimental conditions, as described [10]. Of the 5053 spotted clones, we selected the clones that had detectable expression levels in at least 80% of the experiments (n=2145). Unsupervised hierarchical clustering investigated relationships between samples and relationships between genes. It was applied to data log-transformed and median-centred using the Cluster and TreeView programs (average linkage clustering using Pearson's uncentered correlation as similarity metric) [11]. Using supervised SVM (Support Vector Machine) method, we

investigated the discrimination between CM-R and CM-S mice based on gene expression profiles [12].

Statistical analyses were performed using the TIGR MeV (MultiExperiment Viewer) v3.1 software (http://www.tm4.org/mev.html). Figure 1 shows an outline of data analysis. A one-way ANOVA and SAM (Significant Analysis of Microarrays) procedures [13] were applied to look for strain- and CM-R/CM-S-specific variation in gene expression in the full data set. Welch t-statistics was used to compare CM-R and CM-S mice [14]. Multiple test corrections were performed [15].

Gene annotation of all 2145 analysable genes/ESTs was obtained using EASE (Expression Analysis Systematic Explorer) software [16]. This program which allows a biological interpretation of gene clusters on the basis of GO terms was used to assess whether specific biological pathways were over-represented among the differentially expressed genes and within specific gene clusters. A score based on Fisher Exact test reports the probability that the prevalence of a particular theme within a cluster is due to chance alone given the prevalence of that theme in the population of all genes under study.

All data are MIAME compliant and have been submitted to ArrayExpress database (www.ebi.ac.uk/arrayexpress).

Results

To identify genes differentially expressed between CM-S and CM-R mice, we combined unsupervised and supervised clustering methods with several statistical approaches (figure 1). All the CBA/J mice (*n*=4) and C57BL/6 mice (*n*=5) developed CM while the BALB/c mice (*n*=5) and DBA/2 mice (*n*=5) did not, in spite of similar levels of parasitemia (10.4% ±6.4). The mRNA from the 19 different brain mouse samples was hybridized with cDNA arrays. The mRNA samples from each strain represented biological replicates. The degree of correlation between

gene expression levels of two samples from each strain was high (Pearson's correlation coefficient R mean=0.89).

To analyze the overall expression patterns, we used an unsupervised hierarchical clustering method that groups genes on the vertical axis and samples on the horizontal axis on the basis of similarity in their expression profiles (figure 2*A*). The similarities are summarized in a dendrogram, in which the pattern and length of the branches reflects the relatedness of the samples (figure 2*B*). On this basis, 8 of 10 CM-R mice and 9 of 9 CM-S mice were clustered together, respectively. The discrimination between CM-R and CM-S mice was highly significant (Fisher exact test : $P<.0001$). We assumed that 2 CM-R mice were incorrectly classified because of confounding genes, which were not differentially expressed between mouse groups.

We, therefore, performed one-way ANOVA to look for strain- and CM-R/CM-S-specific variation in gene expression in the full data set (figure 1). We generated empirical *P* values for each gene. Then, we selected all genes having $P <.05$, and we defined a set of 292 informative genes**.** We performed further unsupervised hierarchical clustering of samples based on the expression of the selected genes (figure 2*C*). As shown in figure 2*D*, this procedure successfully classified 10/10 CM-R mice and 9/9 CM-S mice. All mice of each strain were also correctly classified.

To ascertain the accuracy of the prediction of CM molecular signature, we used a "one-out-iterative cross-validation" testing procedure developed by the supervised SVM method. Iteratively, one of the 19 mouse samples was removed from the group, and was classified on the basis of the correlation between its expression profile and the median profile of samples from CM-R mice and CM-S mice. The prediction was considered correct if it corresponded to the actual group. Using this approach, we correctly classified 10 of 10 CM-R mice and 9 of 9 CM-S mice, and we confirmed the results of unsupervised hierarchical clustering.

To search for significant genes based on differential expression between mouse groups, we carried out more stringent statistical methods, which control the percentage of genes identified by chance. We applied two SAM procedures on the full data set to pick out mouse strain-specific and CM-R/CM-S-specific genes (figure 1). Differentially expressed genes were selected by using a false discovery rate (FDR) of 5%. Multi-class SAM yielded a list of 50 genes that showed a differential expression between mouse strains. Two-class SAM yielded a list of 40 genes that appeared to be differentially expressed between CM-R and CM-S mice. Twenty-one genes belonged to both lists, indicating that the effect of both strain and CM phenotype was detected. We, therefore, found 69 significant genes, which belong to the preliminary list of 292 informative genes.

Unsupervised hierarchical clustering of samples based on the expression of the 69 significant genes successfully classified all the BALB/c, DBA/2, C57BL/6, and CBA/J mice on the one hand, and all the CM-R and CM-S mice on the other hand (figure 3). In the same way, 10 of 10 CM-R mice and 9 of 9 CM-S mice were correctly classified by using supervised SVM method. We observed 7 clusters that were either strain-specific or CM-R/CM-S-specific. Clusters A and E grouped DBA/2 and CBA/J strains vs BALB/c and C57BL/6 strains. Clusters C and G discriminated C57BL/6 and BALB/c mice, respectively. Clusters B, D, and F separated CM-R and CM-S mice.

We further analyzed the genes that we selected as CM-R/CM-S-specific by using the two-class SAM (figure 1). The 40 genes were assessed for bias due to strain-specific variation in gene expression. First, using Welch t-test, we searched for genes that were differentially expressed in brain samples from BALB/c and DBA/2 CM-R strains on the one hand, and C57BL/6 and CBA/J CM-S strains on the other hand. Then, we took into account these differences when we analyzed differential expression between CM-R and CM-S strains.

Among the 40 genes selected as CM-R/CM-S-specific, we found 28 genes, the expression of which was not different between BALB/c and DBA/2 CM-R strains nor between C57BL/6 and

CBA/J CM-S strains (table 1). We showed by using Welch t-test and a FDR of 5% that the expression of 11 genes was lower in CM-S mice than in CM-R mice (figure 3, cluster F), and that the expression of 17 genes was higher in CM-S mice than in CM-R mice (figure 3, clusters B and D).

Among the 40 genes selected as CM-R/CM-S-specific, we found 12 genes, the expression of which was different either between BALB/c and DBA/2 CM-R mice or between C57BL/6 and CBA/J CM-S mice (table 2). Using Welch t-test and a FDR of 5%, we found that the expression of *Icsbp1* and *Vcam1* was lower in CM-R strains than in C57BL/6 and CBA/J CM-S strains, and that the expression of *Ccl27* was higher in CM-S strains than in BALB/c and DBA/2 CM-R strains. The expression of the other genes did not significantly differ between CM-R and CM-S (table 2). In all, we identified 31 genes associated with CM.

The analysis of the GO terms of genes composing the different identified clusters by the EASE software showed an overrepresentation of genes involved in defence response (cluster D, $P<.0004$; cluster F, $P<.008$) and immune response (cluster D, $P<.0002$), compared to the population of all genes under study. For instance, the cluster D contains *Samhd1*, *Serping1*, *Icsbp1*, *C1qb*, *Ifit3* and *Ccl27*, which are annotated with the "immune response" GO term.

By using immunochemistry, we further analyzed the expression of two genes associated with susceptibility to CM. We chose *Icsbp1*, which is a crucial component of the IFNγ response, and *Vcam1,* which is a *TNF*-induced gene. The anti-ICSBP1 specific immunostaining, however, was not efficient enough to perform quantification. In contrast, the results were satisfactory for anti-VCAM-1 antibody (figure 4), and we found that *Vcam1* was significantly overexpressed in CM-S mice as compared to CM-R ($P<.0001$, Student t-test).

Discussion

By combining transcriptional profiling with rigorous statistical methods and cluster analysis, we searched for differences in brain gene expression between CM-S (C57BL/6 and CBA/J) and CM-R (BALB/c and DBA/2) mice. We used 5 biological replicates per group to take into account variability between mice. This allowed us to search for genes, which show small but significant changes in expression, and which may be biologically important. The approach used by others was rather to perform one or two microarrays for each condition or to pool the samples in order to minimize the experimental variability. Moreover, the selection of genes is generally based on non-statistically motivated criteria, such as « 2-fold-change » [17-19]. On the basis of a FDR of 5%, we identified a set of 69 genes, the expression of which was significantly different between mouse strains and/or between CM-R and CM-S mice. The cluster analysis showed that this set of genes fully discriminated between mouse strains and between CM-R and CM-S mice.

As several mouse malaria resistance loci have been mapped, we looked at the chromosomal localization of the selected genes. From the 69 selected genes, 67 had a reliable chromosomal location, the 2 missing data corresponding to 2 ESTs. Among the 67 genes, several were located at chromosomal regions previously shown to contain susceptibility loci for cerebral malaria caused by PbA. *Usf1* and *Pafah1b1* were located at the chromosome 1 region and the chromosome 11 region, which contain Berr1 and Berr2, respectively [20]. In addition, *Nfatc1* and *Egr1* were located at chromosome 18 that harbours the locus Esmr also controlling resistance to mouse CM [21]. We also found that *Igf2r, Fkbp5, Lst1,* and *H2-D1* were located at the chromosome 17 region harbouring the Cmcs locus [22]. Interestingly, the same chromosome 17 region was reported to contain genes controlling resistance to *P. chabaudi* (Char3) [23]. Moreover, three other chromosomal regions containing a locus controlling resistance to *P. chabaudi*, namely Char1 (chromosome 9), Char2 (chromosome 8) and Char4 (chromosome 3) contained *Tcf12, Sntb2*, *Fnta, Icsbp1* and *Vcam1*, which we found differentially expressed between CM-R and CM-S mice [23]. It should be stressed that the chromosomal regions linked to malaria resistance are large. Therefore, microarray analyses should be helpful in identifying

candidate genes lying within those genetic regions and showing a differential expression between CM-R and CM-S mice. Cis-acting polymorphisms may affect both gene expression levels and CM resistance.

We identified a cluster of genes up-regulated in C57BL/6 CM-S strain, compared to the CM-R strains and also the CBA CM-S strain. This is consistent with the data indicating that some CM mediators partly differ in CBA and C57BL/6 mice. In particular, it is thought that TNF may be a critical mediator of CM in CBA mice [24], while LTA may be a critical mediator of CM in C57BL/6 mice [25]. In the same way, we identified a cluster of genes specifically up-regulated in BALB/c CM-R mice, suggesting that some CM-R genetic factors may differ between BALB/c and DBA/2 CM-R mice.

We identified three gene clusters associated with resistance or susceptibility to CM. Using Welch t-test and multiple test correction, we further confirmed that the expression of 31 genes significantly differ between CM-R and CM-S mice. This suggests that these genes and related physiological pathways may be critical in malaria pathogenesis in mice. To explore the possible functional role of these genes, we analyzed the occurrence of the GO terms within the clusters. The EASE analysis revealed that the genes clustered preferentially belonged to the two following biological process classes: defence response and immune response. Interestingly, the GO terms are organized in a hierarchical way [26], and the over-represented terms within the clusters of interest were related. Indeed, a number of known genes involved in the immune system and in the pro-inflammatory response had an expression associated with resistance or susceptibility to CM.

In particular, *Gzmb, Igf2r, Ctla2a,* and *C1qb* were found to be up-regulated in CM-S mice. *Gzmb (Ctla1)* encodes granzyme B in cytotoxic T cells and Natural Killer cells [27,28], and induces apoptosis by caspase activation. *Igf2r* has been reported to bind *Gzmb* and to be essential for cytotoxic T cell-mediated apoptosis [29]. This is consistent with the data showing that perforin-deficient C57BL/6 mice are resistant to CM. This indicates that the perforin/granzyme cytolytic pathway may be a major mechanism employed by CD8+ T cells during CM [30,31]. *Ctla2a* is

expressed in mouse activated T cells [32]. *C1qb* that is the complement component 1 beta gene has also been shown to be up-regulated after permanent focal ischemia in the mouse [33]. In addition, overexpression of cyclooxygenase-2 (COX-2), which is expressed in mouse CM brain [34], leads to selective induction of *C1qb* expression in brain [35]. One might expect that the presence of locally produced complement components within brain is likely to induce a potent inflammatory response, and to contribute to brain damage.

Pafhab1, which was down-regulated in CM-S mice, encodes a subunit Platelet Activating Factor acetylhydrolase, which regulates the level of platelet activating factor (PAF) in brain [36]. It has been suggested that PAF could play a pathophysiological role in severe malaria through activating platelets [37]. Platelets that are sequestered in cerebral microvessels are thought to be involved in endothelial cell damage and in malaria pathogenesis [1].

We identified several genes known to be related to IFN or TNF responses. *Samhd1*, *Ifit3* and *Lst1* are interferon-regulated genes [38-40], and *Icsbp1*, *Stat1* are important components of the cellular response to IFNγ. *Nfatc1* binds to *Icsbp1* to promote the transcription of *Il12B* [41], and to *Egr-1* to promote the transcription of *Tnf* and *Il2* [42]. *Vcam-1* and *Ccl27* are TNF-induced genes [43,44], while *Fkbp5* encodes a heat shock protein (Hsp) 90-binding immunophilin implicated in the TNF signal transduction pathway [45]. We found that *Samhd1*, *Ifit3*, *Icsbp1*, *Stat1, Nfatc1, Fkbp5,* and *Vcam-1* expression level was nominally higher in the brain of CM-S mice than in the brain of CM-R mice. We analyzed the expression of *Vcam-1* and *Icsbp1* to validate the DNA microarray data by using immunochemistry method. We were unable to detect *Icsbp1*, indicating that the monoclonal antibody (mAb) did not perform well on cryosections. In contrast, *Vcam-1* results confirmed the correlation between *Vcam-1* expression level and susceptibility to CM. In the same way, Lou et al [46] showed that TNF induced a higher level of *ICAM-1* and *VCAM-1* expression in CBA/J (CM-S) than in BALB/c (CM-R) brain microvascular endothelial cells (MVECs). In addition, Sexton et al [47] recently showed that interferon-regulated gene transcripts or IFN regulatory factors, such as *Stat1*, strongly increased

in brain from C57BL/6 (CM-S) mice infected by PbA. These results are in line with the current knowledge of the physiological pathways involved in malaria pathogenesis. The role of IFNγ, TNF and their receptors in mouse CM has been demonstrated by using neutralizing mAb [48] or gene targeting [49-51]. Moreover, human TNF and IFNγR1 polymorphisms are associated with susceptibility to CM [52,53]. High levels of TNF and IFNγ have been measured in patients with severe malaria [48,54], and TNFR2 complex has been observed on the surface of MVECs in human CM [1] . NK cells, CD4+ T cells and γδ T cells were shown to produce IFNγ in in vitro models of *P. falciparum* infection [55-57]. In mice, NK T cells or NK cells could be important sources of IFNγ, since the introduction of C57BL/6 (CM-S) strain NKC genes into a BALB/c (CM-R) strain results in an increased production of IFNγ, an increased Th1 response, and susceptibility to CM [58].

Here, we show that there are gene expression profiles associated with susceptibility to mouse CM. We identified several genes, the expression of which significantly differs between CM-S and CM-R mice. The concordance between our data and those from literature with respect to the function of genes strengthens the validity of our results. This suggests that these genes -or genes involved in the same physiological pathway- are implicated in malaria pathogenesis in mice and can be considered as candidate genes in humans. Microarray analyses should provide new insights into the knowledge of key genes governing the response to malaria.

Acknowledgments

We thank Béatrice Loriod and Geneniève Victorero for technical support and advices.

References

1. Hunt NH, Grau GE. Cytokines: accelerators and brakes in the pathogenesis of cerebral malaria. Trends Immunol 2003;24:491-9

2. De Souza JB, Riley EM. Cerebral malaria: the contribution of studies in animal models to our understanding of immunopathogenesis. Microbes Infect 2002;4:291-300

3. Rihet P, Traore Y, Abel L, Aucan C, Traore-Leroux T and Fumoux F. Malaria in humans: Plasmodium falciparum blood infection levels are linked to chromosome 5q31-q33. Am J Hum Genet 1998;63:498-505

4. Garcia A, Marquet S, Bucheton B, et al. Linkage analysis of blood Plasmodium falciparum levels: interest of the 5q31-q33 chromosome region. Am J Trop Med Hyg 1998;58:705-9

5. Jepson A, Sisay-Joof F, Banya W, et al. Genetic linkage of mild malaria to the major histocompatibility complex in Gambian children: study of affected sibling pairs. Bmj 1997;315:96-7

6. Flori L, Sawadogo S, Esnault C, Delahaye NF, Fumoux F and Rihet P. Linkage of mild malaria to the major histocompatibility complex in families living in Burkina Faso. Hum Mol Genet 2003;12:375-8

7. Lou J, Lucas R and Grau GE. Pathogenesis of cerebral malaria: recent experimental data and possible applications for humans. Clin Microbiol Rev 2001;14:810-20, table of contents

8. Combes V, Coltel N, Alibert M, et al. ABCA1 gene deletion protects against cerebral malaria: potential pathogenic role of microparticles in neuropathology. Am J Pathol 2005;166:295-302

9. Puthier D, Joly F, Irla M, et al. A general survey of thymocyte differentiation by transcriptional analysis of knockout mouse models. J Immunol 2004;173:6109-18

10. Bertucci F, Salas S, Eysteries S, et al. Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. Oncogene 2004;23:1377-91

11. Eisen MB, Spellman PT, Brown PO and Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863-8

12. Brown MP, Grundy WN, Lin D, et al. Knowledge-based analysis of microarray gene expression data by using support vector machines. Proc Natl Acad Sci U S A 2000;97:262-7

13. Tusher VG, Tibshirani R and Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116-21

14. Welch BL. The generalization of 'students' problem when several different population variances are involved. Biometrika 1947;34:28-35

15. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. J Royal Stat Soc Ser B 1995;57:289-300

16. Hosack DA, Dennis G, Jr., Sherman BT, Lane HC and Lempicki RA. Identifying biological themes within lists of genes with EASE. Genome Biol 2003;4:R70

17. Lemon WJ, Bernert H, Sun H, Wang Y and You M. Identification of candidate lung cancer susceptibility genes in mouse using oligonucleotide arrays. J Med Genet 2002;39:644-55

18. Ylostalo J, Randall AC, Myers TA, Metzger M, Krogstad DJ and Cogswell FB. Transcriptome profiles of host gene expression in a monkey model of human malaria. J Infect Dis 2005;191:400-9

19. Keller C, Lauber J, Blumenthal A, Buer J and Ehlers S. Resistance and susceptibility to tuberculosis analysed at the transcriptome level: lessons from mouse macrophages. Tuberculosis (Edinb) 2004;84:144-58

20. Bagot S, Campino S, Penha-Goncalves C, Pied S, Cazenave PA and Holmberg D. Identification of two cerebral malaria resistance loci using an inbred wild-derived mouse strain. Proc Natl Acad Sci U S A 2002;99:9919-23

21. Nagayasu E, Nagakura K, Akaki M, et al. Association of a determinant on mouse chromosome 18 with experimental severe Plasmodium berghei malaria. Infect Immun 2002;70:512-6

22. Ohno T, Nishimura M. Detection of a new cerebral malaria susceptibility locus, using CBA mice. Immunogenetics 2004;56:675-8

23. Fortin A, Stevenson MM and Gros P. Complex genetic control of susceptibility to malaria in mice. Genes Immun 2002;3:177-86

24. Grau GE, Fajardo LF, Piguet PF, Allet B, Lambert PH and Vassalli P. Tumor necrosis factor (cachectin) as an essential mediator in murine cerebral malaria. Science 1987;237:1210-2

25. Engwerda CR, Mynott TL, Sawhney S, De Souza JB, Bickle QD and Kaye PM. Locally up-regulated lymphotoxin alpha, not systemic tumor necrosis factor alpha, is the principle mediator of murine cerebral malaria. J Exp Med 2002;195:1371-7

26. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25-9

27. Pardo J, Bosque A, Brehm R, et al. Apoptotic pathways are selectively activated by granzyme A and/or granzyme B in CTL-mediated target cell lysis. J Cell Biol 2004;167:457-68

28. Pardo J, Balkow S, Anel A and Simon MM. Granzymes are essential for natural killer cell-mediated and perf-facilitated tumor control. Eur J Immunol 2002;32:2881-7

29. Motyka B, Korbutt G, Pinkoski MJ, et al. Mannose 6-phosphate/insulin-like growth factor II receptor is a death receptor for granzyme B during cytotoxic T cell-induced apoptosis. Cell 2000;103:491-500

30. Nitcheu J, Bonduelle O, Combadiere C, et al. Perforin-dependent brain-infiltrating cytotoxic CD8+ T lymphocytes mediate experimental cerebral malaria pathogenesis. J Immunol 2003;170:2221-8

31. Potter S, Chaudhri G, Hansen A and Hunt NH. Fas and perforin contribute to the pathogenesis of murine cerebral malaria. Redox Rep 1999;4:333-5

32. Denizot F, Brunet JF, Roustan P, et al. Novel structures CTLA-2 alpha and CTLA-2 beta expressed in mouse activated T cells and mast cells and homologous to cysteine proteinase proregions. Eur J Immunol 1989;19:631-5

33. Van Beek J, Chan P, Bernaudin M, Petit E, MacKenzie ET and Fontaine M. Glial responses, clusterin, and complement in permanent focal cerebral ischemia in the mouse. Glia 2000;31:39-50

34. Ball HJ, MacDougall HG, McGregor IS and Hunt NH. Cyclooxygenase-2 in the pathogenesis of murine cerebral malaria. J Infect Dis 2004;189:751-8

35. Spielman L, Winger D, Ho L, Aisen PS, Shohami E and Pasinetti GM. Induction of the complement component C1qB in brain of transgenic mice with neuronal overexpression of human cyclooxygenase-2. Acta Neuropathol (Berl) 2002;103:157-62

36. Hattori M, Aoki J, Arai H and Inoue K. PAF and PAF acetylhydrolase in the nervous system. J Lipid Mediat Cell Signal 1996;14:99-102

37. Davis TM, Sturm M, Zhang YR, et al. Platelet-activating factor and lipid metabolism in acute malaria. J Infect 1993;26:279-85

38. Lafuse WP, Brown D, Castle L and Zwilling BS. Cloning and characterization of a novel cDNA that is IFN-gamma-induced in mouse peritoneal macrophages and encodes a putative GTP-binding protein. J Leukoc Biol 1995;57:477-83

39. Smith JB, Herschman HR. The glucocorticoid attenuated response genes GARG-16, GARG-39, and GARG-49/IRG2 encode inducible proteins containing multiple tetratricopeptide repeat domains. Arch Biochem Biophys 1996;330:290-300

40. de Baey A, Fellerhoff B, Maier S, Martinozzi S, Weidle U and Weiss EH. Complex expression pattern of the TNF region gene LST1 through differential regulation, initiation, and alternative splicing. Genomics 1997;45:591-600

41. Zhu C, Rao K, Xiong H, et al. Activation of the murine interleukin-12 p40 promoter by functional interactions between NFAT and ICSBP. J Biol Chem 2003;278:39372-82

42. Decker EL, Nehmann N, Kampen E, Eibel H, Zipfel PF and Skerka C. Early growth response proteins (EGR) and nuclear factors of activated T cells (NFAT) form heterodimers and regulate proinflammatory cytokine gene expression. Nucleic Acids Res 2003;31:911-21

43. Neish AS, Williams AJ, Palmer HJ, Whitley MZ and Collins T. Functional analysis of the human vascular cell adhesion molecule 1 promoter. J Exp Med 1992;176:1583-93

44. Vestergaard C, Johansen C, Otkjaer K, Deleuran M and Iversen L. Tumor necrosis factor-alpha-induced CTACK/CCL27 (cutaneous T-cell-attracting chemokine) production in keratinocytes is controlled by nuclear factor kappaB. Cytokine 2005;29:49-55

45. Bouwmeester T, Bauch A, Ruffner H, et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. Nat Cell Biol 2004;6:97-105

46. Lou J, Gasche Y, Zheng L, et al. Differential reactivity of brain microvascular endothelial cells to TNF reflects the genetic susceptibility to cerebral malaria. Eur J Immunol 1998;28:3989-4000

47. Sexton AC, Good RT, Hansen DS, et al. Transcriptional profiling reveals suppressed erythropoiesis, up-regulated glycolysis, and interferon-associated responses in murine malaria. J Infect Dis 2004;189:1245-56

48. Grau GE, Piguet PF, Vassalli P and Lambert PH. Tumor-necrosis factor and other cytokines in cerebral malaria: experimental and clinical data. Immunol Rev 1989;112:49-70

49. Yanez DM, Manning DD, Cooley AJ, Weidanz WP and van der Heyde HC. Participation of lymphocyte subpopulations in the pathogenesis of experimental murine cerebral malaria. J Immunol 1996;157:1620-4

50. Amani V, Vigario AM, Belnoue E, et al. Involvement of IFN-gamma receptor-medicated signaling in pathology and anti-malarial immunity induced by Plasmodium berghei infection. Eur J Immunol 2000;30:1646-55

51. Lucas R, Juillard P, Decoster E, et al. Crucial role of tumor necrosis factor (TNF) receptor 2 and membrane-bound TNF in experimental cerebral malaria. Eur J Immunol 1997;27:1719-25

52. Knight JC, Udalova I, Hill AV, et al. A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. Nat Genet 1999;22:145-50

53. Koch O, Awomoyi A, Usen S, et al. IFNGR1 gene promoter polymorphisms and susceptibility to cerebral malaria. J Infect Dis 2002;185:1684-7

54. Ho M, Sexton MM, Tongtawe P, Looareesuwan S, Suntharasamai P and Webster HK. Interleukin-10 inhibits tumor necrosis factor production but not antigen-specific lymphoproliferation in acute Plasmodium falciparum malaria. J Infect Dis 1995;172:838-44

55. Goodier MR, Lundqvist C, Hammarstrom ML, Troye-Blomberg M and Langhorne J. Cytokine profiles for human V gamma 9+ T cells stimulated by Plasmodium falciparum. Parasite Immunol 1995;17:413-23

56. Artavanis-Tsakonas K, Riley EM. Innate immune response to malaria: rapid induction of IFN-gamma from human NK cells by live Plasmodium falciparum-infected erythrocytes. J Immunol 2002;169:2956-63

57. Scragg IG, Hensmann M, Bate CA and Kwiatkowski D. Early cytokine induction by Plasmodium falciparum is not a classical endotoxin-like process. Eur J Immunol 1999;29:2636-44

58. Hansen DS, Evans KJ, D'Ombrain MC, et al. The natural killer complex regulates severe malarial pathogenesis and influences acquired immune responses to Plasmodium berghei ANKA. Infect Immun 2005;73:2288-97

Figure legends

Figure 1.

A schematic outline of step-wise data analysis. SAM, Significant Analysis of Microarrays. FDR, False Discovery Rate. HCL, Hierarchical Clustering. SVM, Support Vector Machine. CM, cerebral malaria. CM-R, CM-resistant mice. CM-S, CM-susceptible mice. A one-way ANOVA was applied to look for strain- and CM-R/CM-S-specific variation in gene expression in the full data set. Empirical *P*-values were computed for each gene by bootstrapping from 10,000 permutations. All genes having $P<.05$ were selected for unsupervised and supervised clusterings. To estimate the accuracy of prediction of the "CM molecular signature", we applied the "one-out-iterative cross-validation" testing procedure developed by the supervised SVM method. Briefly, the algorithm withholds one of the 19 mice samples, builds a training set based only on the remaining samples, and predicts the class of the withheld sample. The process was repeated for each sample, and a Fisher exact test was used to assess the significance of the classification. From the full data set, significant differential gene expression was determined using SAM method. The SAM algorithm gives an estimate of the FDR, which is the proportion of false positives among all of the genes initially identified as being differentially expressed. The multi-class and two-class unpaired SAM procedures were used to identify mouse strain-specific and CM-R/CM-S-specific genes, respectively, on the basis of 10,000 bootstrapped permutations. Missing values were imputed using a K-nearest-neighbor algorithm. Multiple testing problem was addressed by choosing a FDR of 5%. The significant genes with differential expression between CM-R and CM-S mice were extracted from the SAM results with Welch t-test for small samples with unequal variances in the two groups. Empirical *P*-values were computed for each gene from 10,000 bootstrapped permutations. All genes having $P<.05$ were selected and we applied a FDR of 5%.

Figure 2.

Gene expression profiles in CM-R mice and CM-S mice. *A*, Hierarchical clustering of 19 brain tissue samples representing the four mouse strains (BALB/c, DBA/2, C57BL/6, CBA/J) and 2145 cDNA clones (full data set) based on mRNA expression levels. Each row represents a gene and each column represents a sample. The normalized expression value for each gene was depicted according to the colour scale at the bottom. Red and green indicate expression levels above and below the median, respectively. Grey indicates missing data. Dendograms of samples (above matrix) and genes (to the left of matrix) represent overall similarities in gene expression profiles. *B*, Dendogram of samples representing the results of the same global hierarchical clustering applied to the 19 brain tissue samples. Two groups of samples were defined: CM-R mice (BALB/c, DBA/2) in blue and CM-S mice (C57BL/6, CBA/J) in red. C, Hierarchical clustering of the 19 brain tissue samples using expression levels of 292 informative genes differentially expressed between the four mouse strains. This group of informative genes was extracted from the full data set (n=2145) by an one-way ANOVA test on the basis of 10,000 permutations (P<.05). *D*, Dendogram of the samples representing the results of the same hierarchical clustering applied to the 292 informative genes.

Figure 3.

Hierarchical classification of strain-specific and CM-R and CM-S significant genes. Hierarchical clustering of the 19 brain tissue samples using expression levels of 69 significant genes differentially expressed between the four mouse strains (BALB/c, DBA/2, C57BL/6, CBA/J). This group of significant genes was extracted from the full data set (n=2145) by a SAM procedure on the basis of 10,000 permutations, and a FDR of 5%. The Clone ID and the UniGene Symbol are indicated for each gene. Coloured bars to the right indicate the location of clusters (A-G).

Figure 4.

***VCAM-1* expression in brain vessels of CM-S and CM-R mice.** *A*, Histological pictures of CM-S and *B*, CM-R mice. *VCAM-1* expression was expressed for each vessel as a ratio of the red chromogen area out of the total vessel lumen. *C*, *VCAM-1* was significantly overexpressed in CM-S mice (Student t-test).

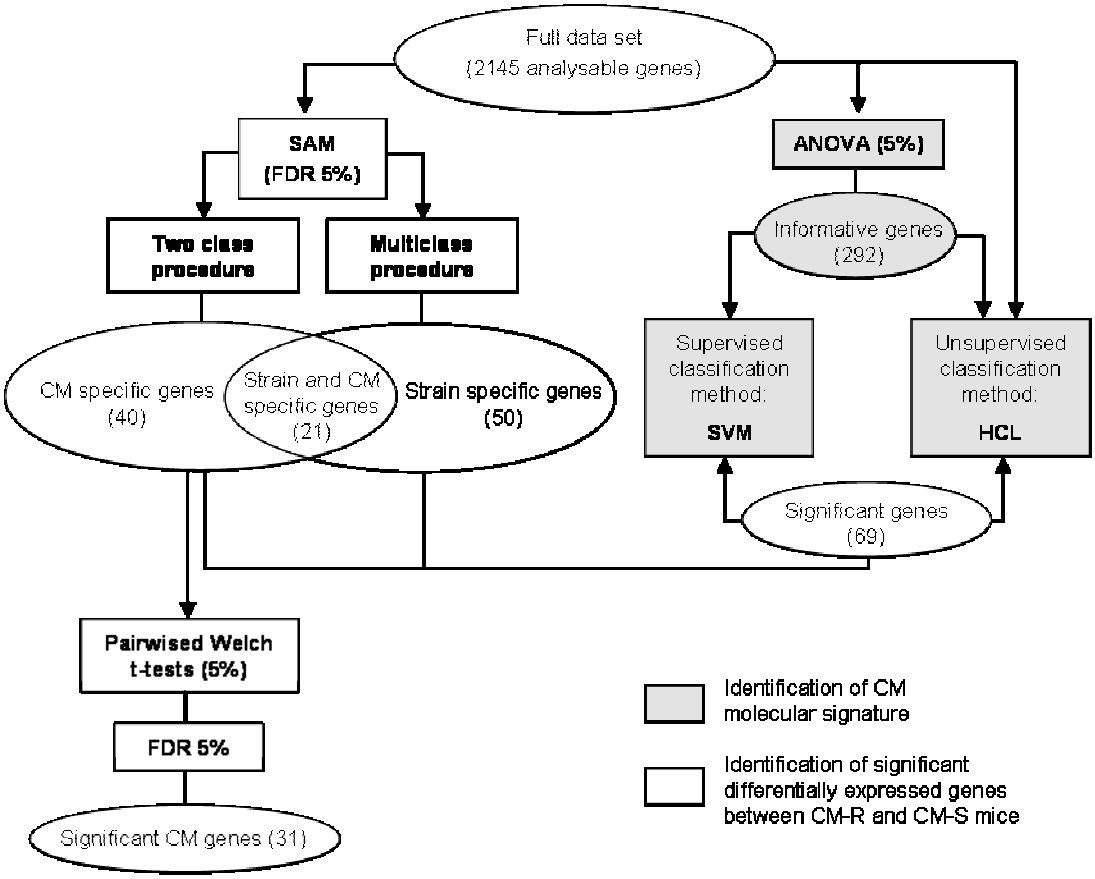

Figure 1.

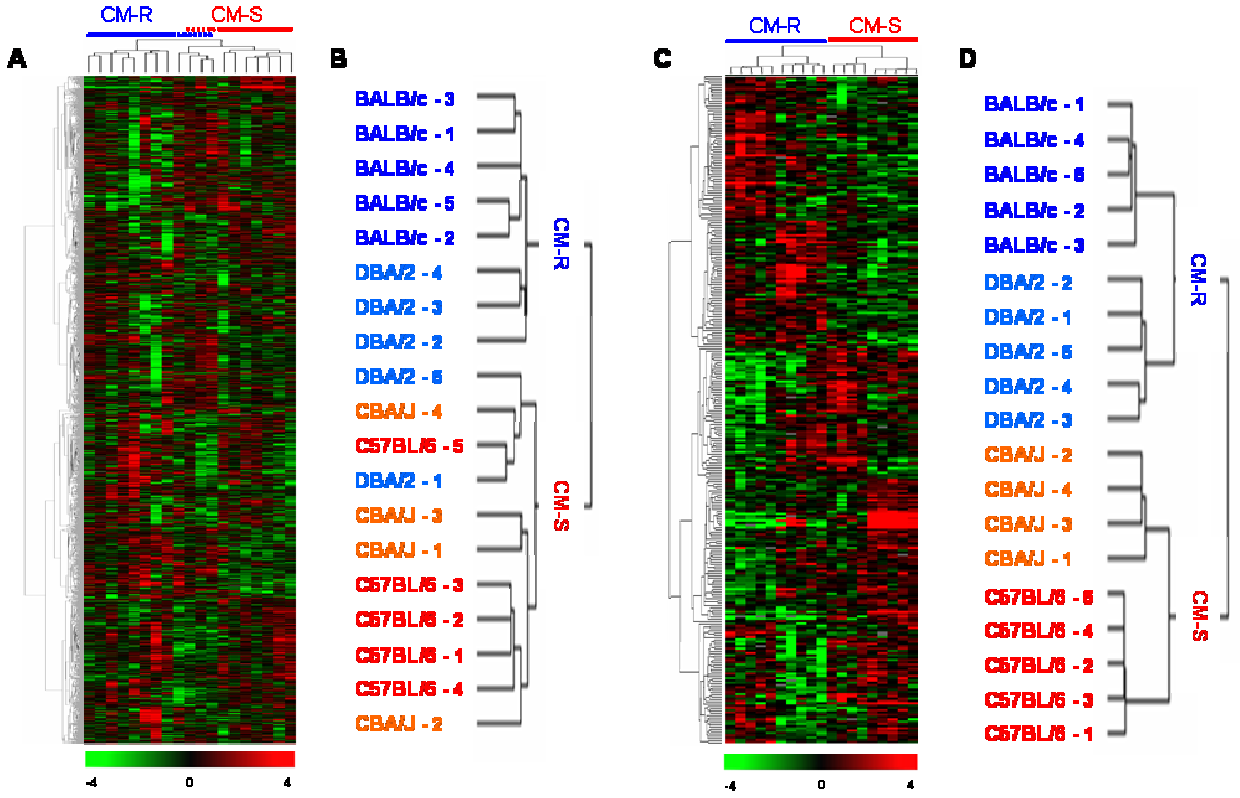

Figure 2.

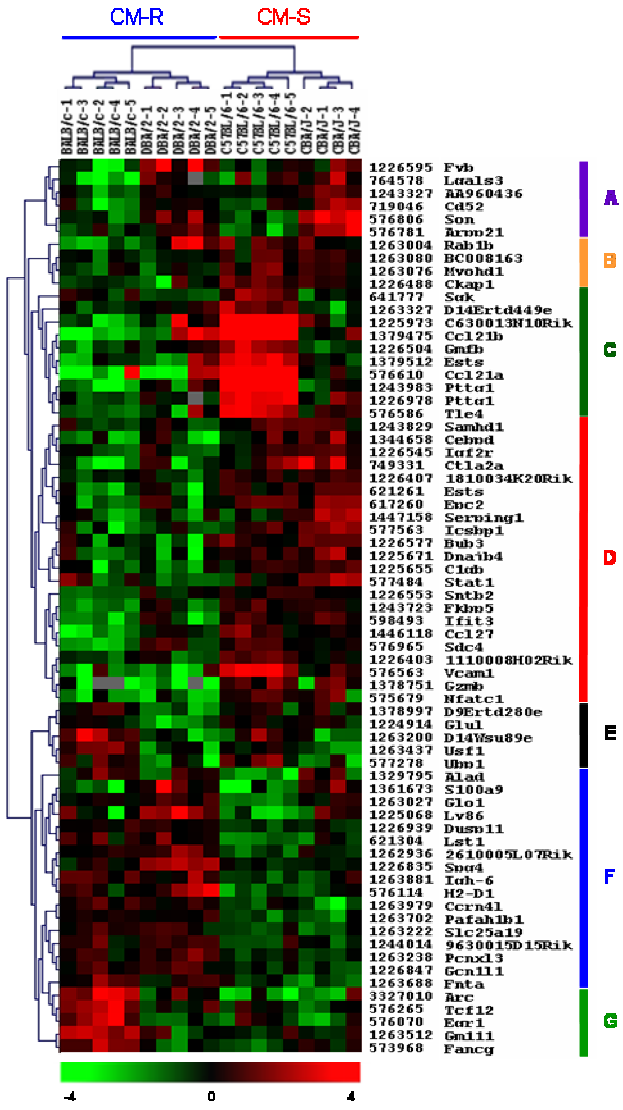

Figure 3.

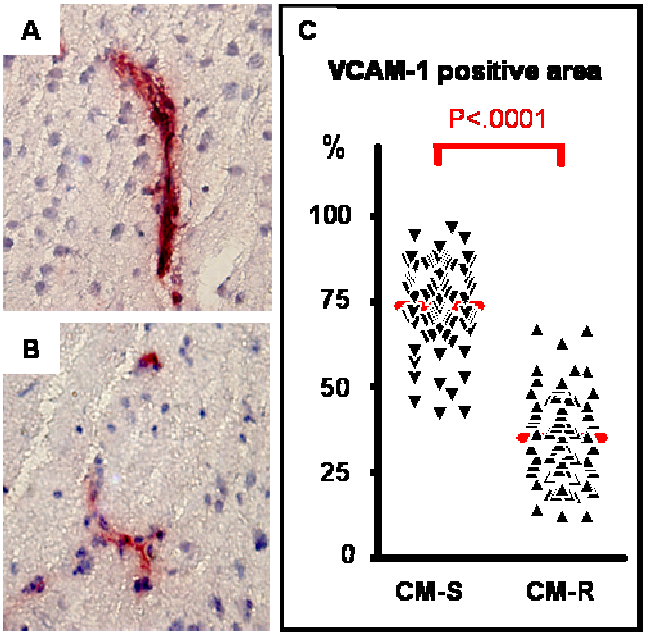

Figure 4.

Table 1. Cerebral malaria (CM) specific genes obtained from Welch t-test and FDR 5%.

Gene symbol (Cluster)	Name	CloneID	MGI ID	Score[a]	Empirical P value
BC008163 (B)	CDNA sequence BC008163	1263080	2385211	3.67	.0018
Myohd1 (B)	Myosin head domain containing 1	1263076	1913446	3.72	.0017
Ckap1 (B)	Cytoskeleton-associated protein 1	1226488	1913661	3.51	.0009
Samhd1 (D)	SAM domain and HD domain. 1	1243829	1927468	3.60	.0042
Cebpd (D)	Transcribed locus // CCAAT/enhancer binding protein (C/EBP). delta	1344658	103573	3.40	.0039
Igf2r (D)	Insulin-like growth factor 2 receptor	1226545	96435	3.73	.0029
Ctla2a (D)	Cytotoxic T lymphocyte-associated protein 2 alpha	749331	88554	3.89	.0013
1810034K20Rik (D)	RIKEN cDNA 1810034K20 gene	1226407	1915131	4.72	<.00001
Ests (D)	-	621261	-	3.79	.0009
Epc2 (D)	Enhancer of polycomb homolog 2 (Drosophila)	617260	1278321	4.04	.0007
Bub3 (D)	Budding uninhibited by benzimidazoles 3 homolog (S. cerevisiae)	1226577	1343463	3.31	.0016
Dnajb4 (D)	DnaJ (Hsp40) homolog. subfamily B. member 4	1225671	1914285	3.68	.0016
C1qb (D)	Complement component 1. q subcomponent. beta polypeptide	1225655	88224	3.74	.0006
Fkbp5 (D)	FK506 binding protein 5	1243723	104670	4.69	.0006
Ifit3 (D)	Interferon-induced protein with tetratricopeptide repeats 3	598493	1101055	4.46	.0002
Gzmb (D)	Granzyme B	1378751	109267	4.44	.0009
Nfatc1 (D)	Nuclear factor of activated T-cells. cytoplasmic. calcineurin-dependent 1	575679	102469	4.37	.0013
Dusp11 (F)	Dual specificity phosphatase 11 (RNA/RNP complex 1-interacting)	1226939	1919352	-4.87	.0001
Lst1 (F)	Leukocyte specific transcript 1	621304	1096324	-4.37	.0008
Igh-6 (F)	Immunoglobulin heavy chain 6 (heavy chain of IgM)	1263881	96448	-4.03	.0011
H2-D1 (F)	Histocompatibility 2. D region locus 1	576114	95896	-3.83	.0013
Ccrn4l (F)	CCR4 carbon catabolite repression 4-like (S. cerevisiae)	1263979	109382	-3.54	.0004
Pafah1b1 (F)	Platelet-activating factor acetylhydrolase. isoform 1b. beta1 subunit	1263702	109520	-4.79	<.00001
Slc25a19 (F)	Solute carrier family 25 (mitochondrial deoxynucleotide carrier). member 19	1263222	1914533	-4.66	.0004
9630015D15Rik (F)	RIKEN cDNA 9630015D15 gene	1244014	2140359	-3.67	.0036
Pcnxl3 (F)	Pecanex-like 3 (Drosophila)	1263238	1861733	-4.18	.0013
Gcn1l1 (F)	GCN1 general control of amino-acid synthesis 1-like 1 (yeast)	1226847	2444248	-4.60	.0004
Fnta (F)	Farnesyltransferase. CAAX box. alpha	1263688	104683	-3.92	.0007

NOTE. The genes are listed following the clustering order in figure 2. CM-R, CM-resistant mice. CM-S, CM-susceptible mice.

[a] Positive score: CM-S mice gene expression was higher than CM-R mice gene expression.

Negative Score: CM-R mice gene expression was higher than CM-S mice gene expression.

Table 2. Cerebral malaria (CM) specific genes with a strain effect obtained from Welch t-test and FDR 5%.

Group[a], Gene symbol (Cluster)	Name	CloneID	MGI ID	Welch t-test	Score[b]	Empirical P value[c]
Group A						
Ccl21a (C)	Chemokine (C-C motif) ligand 21a (leucine)	576610	1349182	CM-R vs C57BL/6	8.18	<.00001*
				CM-R vs CBA/J	2.43	…
Tle4 (C)	Transducin-like enhancer of split 4, homolog of Drosophila E(spl)	576586	104633	CM-R vs C57BL/6	6.73	.00033*
				CM-R vs CBA/J	2.02	…
Icsbp1 (D)	Interferon consensus binding protein 1	577563	96395	CM-R vs C57BL/6	2.47	.033*
				CM-R vs CBA/J	6.36	<.00001*
1110008H02Rik (D)	RIKEN cDNA 1110008H02 gene	1226403	1921074	CM-R vs C57BL/6	6.58	<.00001*
				CM-R vs CBA/J	2.31	...
Vcam1 (D)	Vascular cell adhesion molecule 1	576563	98926	CM-R vs C57BL/6	6.29	<.00001*
				CM-R vs CBA/J	2.92	.016*
Group B						
Ccl27 (D)	Chemokine (C-C motif) ligand 27	1446118	1343459	BALB/c vs CM-S	6.31	<.00001*
				DBA/2 vs CM-S	2.40	.032*
Arc (G)	Activity regulated cytoskeletal-associated protein	3327010	88067	BALB/c vs CM-S	-5.73	.0005*
				DBA/2 vs CM-S	-2.50	.037
Group C						
C630013N10Rik (C)	RIKEN cDNA C630013N10 gene	1225973	2443079	BALB/c vs C57BL/6	10.30	<.00001*
				BALB/c vs CBA/J	2.22	…
				DBA/2 vs C57BL/6	4.28	<.00001*
				DBA/2 vs CBA/J	0.41	…
Ccl21b (C)	Chemokine (C-C motif) ligand 21b (serine)	1379475	1349183	BALB/c vs C57BL/6	11.74	<.00001*
				BALB/c vs CBA/J	3.84	<.00001*
				DBA/2 vs C57BL/6	5.34	<.00001*
				DBA/2 vs CBA/J	0.73	…
Pttg1 (C)	Pituitary tumor-transforming 1	1243983[d] (1226978)	1353578	BALB/c vs C57BL/6	15.05	<.00001*
				BALB/c vs CBA/J	2.66	.031*
				DBA/2 vs C57BL/6	10.32	<.00001*
				DBA/2 vs CBA/J	0.77	…
2610005L07Rik (F)	RIKEN cDNA 2610005L07 gene	1262936	1914283	BALB/c vs C57BL/6	-4.53	<.00001*
				BALB/c vs CBA/J	-1.63	…
				DBA/2 vs C57BL/6	-4.81	<.00001*
				DBA/2 vs CBA/J	-2.90	.023*

NOTE. The genes, the expression of which significantly differed between CM-R and CM-S mice, are shown **in bold**. A false discovery rate (FDR) of 5% was used.

[a] Group A: the expression of the gene differed between C57BL/6 and CBA/J mice.

Group B: the expression of the gene differed between BALB/c and DBA/2 mice.

Group C: the expression of the gene differed between C57BL/6 and CBA/J mice, and between BALB/c and DBA/2 mice.

[b] Positive score: CM-S mice gene expression was higher than CM-R mice gene expression.

Negative Score: CM-R mice gene expression was higher than CM-S mice gene expression.

[c] Significant results with a FDR of 5% are indicated by an asterisk.

[d] Results are shown for this clone. Similar results were obtained for the other clone.